C L A I M S

1. A sentence classification device characterized by comprising:

a term list having a plurality of terms each comprising not less than one word;

DT matrix generation means for generating a DT matrix two-dimensionally expressing a relationship between each document contained in a document set and said each term;

DT matrix transformation means for generating a transformed DT matrix having clusters having blocks of associated documents by transforming the DT matrix obtained by said DT matrix generation means on the basis of a DM decomposition method used in a graph theory; and

classification generation means for generating classifications associated with the document set on the basis of a relationship between each cluster on the transformed DT matrix obtained by said DT matrix transformation means and said each document classified according to the clusters.

2. A sentence classification device according to claim 1, characterized in that said classification generation means comprises document classification means for outputting, for each cluster on the transformed DT matrix obtained by said DT matrix transformation means, documents belonging to the cluster as the same classification.

3. A sentence classification device according to claim 2, characterized by further comprising label generation means for outputting each term strongly connected to each document belonging to said arbitrary cluster as a label indicating a classification of the cluster.

4. A sentence classification device according to claim 2, characterized by further comprising document organization means for sequentially outputting documents belonging to said arbitrary cluster or all documents in an arrangement order of the documents in the transformed DT matrix.

5. A sentence classification device according to claim 2, characterized by further comprising summary generation means for outputting, as a summary of said arbitrary document, a sentence of sentences constituting the document which contains a term strongly connected to the document.

6. A sentence classification device according to claim 2, characterized by further comprising:

term list edition means for adding or deleting an arbitrary term with respect to the term list; and

index generation means for making said DT matrix generation means generate DT matrices by using term lists before and after edition by said term list edition means, and generating and outputting an index indicating validity of the edition from the DT matrices.

7. A sentence classification device according to claim 1, characterized in that said classification generation means comprises:

virtual representative document generation means for generating a virtual representative document, for each cluster on a transformed DT matrix, from a term of each document belonging to the cluster; and

large classification generation means for generating a large classification of documents by repeatedly performing clustering processing of setting a DT matrix generated by said DT matrix generation means in an initial state, causing said virtual representative document generation means to generate a virtual representative document for each cluster on a transformed DT matrix generated from the DT matrix by said DT matrix transformation means, generating a new DT matrix used for next clustering processing by adding the virtual representative document to the transformed DT matrix and deleting documents belonging to the cluster of the virtual representative document from the transformed DT matrix, and outputting, for said each cluster, information associated with the documents constituting the cluster as large classification data.

8. A sentence classification device according to claim 7, characterized in that said large classification generation means terminates repetition of the clustering processing when no cluster is obtained

from the transformed DT matrix in the clustering processing.

9. A sentence classification device according to claim 7, characterized by further comprising large classification label generation means for, if a virtual representative document is contained in a given cluster of clusters obtained by the clustering processing, generating a label of the cluster on which the virtual representative document is based from a term strongly connected to the virtual representative document.

10. A sentence classification method characterized by comprising:

the DT matrix generation step of generating a DT matrix two-dimensionally expressing a relationship between each document contained in a document set and each term of a term list having a plurality of terms each comprising not less than one word;

the DT matrix transformation step of generating a transformed DT matrix having clusters having blocks of associated documents by transforming the DT matrix on the basis of a DM decomposition method used in a graph theory; and

the classification generation step of generating classifications associated with the document set on the basis of a relationship between each cluster on the transformed DT matrix and said each document classified according to the clusters.

11. A sentence classification method according to claim 10, characterized in that the classification generation step comprises the document classification step of outputting, for each cluster on the transformed DT matrix, documents belonging to the cluster as the same classification.

12. A sentence classification method according to claim 11, characterized by further comprising the step of outputting each term strongly connected to each document belonging to said arbitrary cluster as a label indicating a classification of the cluster.

13. A sentence classification method according to claim 11, characterized by further comprising the step of sequentially outputting documents belonging to said arbitrary cluster or all documents in an arrangement order of the documents in the transformed DT matrix.

14. A sentence classification method according to claim 11, characterized by further comprising the step of outputting, as a summary of a document, a sentence of sentences constituting said arbitrary document which contains a term strongly connected to the document.

15. A sentence classification method according to claim 11, characterized by further comprising:

the step of adding or deleting an arbitrary term with respect to the term list; and

the step of generating DT matrices by using term lists before and after edition, and generating and outputting an index indicating validity of the edition from the DT matrices.

16. A sentence classification method according to claim 10, characterized in that the classification generation step comprises:

the virtual representative document generation step of generating a virtual representative document, for each cluster on a transformed DT matrix, from a term of each document belonging to the cluster; and

the large classification generation step of generating a large classification of documents by repeatedly performing clustering processing comprising the step of setting a DT matrix generated in the DT matrix generation step in an initial state, generating a virtual representative document in the virtual representative document generation step for each cluster on a transformed DT matrix generated from the DT matrix in the DT matrix transformation step, the step of generating a new DT matrix used for next clustering processing by adding the virtual representative document to the transformed DT matrix and deleting documents belonging to the cluster of the virtual representative document from the transformed DT matrix, and the step of outputting, for said each cluster, information associated with the documents constituting the cluster

as large classification data.

17. A sentence classification method according to claim 16, characterized in that in the large classification generation step, repetition of the clustering processing is terminated when no cluster is obtained from the transformed DT matrix in the clustering processing.

18. A sentence classification method according to claim 16, characterized by further comprising the large classification label generation step of, if a virtual representative document is contained in a given cluster of clusters obtained by the clustering processing, generating a label of the cluster on which the virtual representative document is based from a term strongly connected to the virtual representative document.